ELSEVIER

Discrete Mathematics 180 (1998) 173–184

DISCRETE MATHEMATICS

A new way of counting the column-convex polyominoes by perimeter

Svjetlan Feretić

Šetalište Joakima Rakovca 17, 51000 Rijeka, Croatia

Received 18 July 1995; received in revised form 20 June 1996; accepted 15 July 1996

Abstract

We introduce a new class of plane figures: the sequences of tailed column-convex polyominoes (for short: stapoes). Let $G(x, y)$ and $I(x, y)$ denote the perimeter generating functions for column-convex polyominoes and stapoes, respectively. It will be clear from the definitions that $G(x, y)$ is a simple fraction of $I(x, y)$. But this latter function can be DSV-computed by solving just one quadratic equation (and not a system of quadratic equations). Thus the formula for $G(x, y)$ can be obtained with ease.

Résumé

Nous introduisons une nouvelle classe de figures planaires: les chaînes de polyominos verticalement convexes. Soient $G(x, y)$ et $I(x, y)$ les séries génératrices selon le périmètre des polyominos verticalement convexes et des chaînes de polyominos verticalement convexes respectivement. Il sera clair, d'après les définitions, que $G(x, y)$ est une fraction rationnelle simple de $I(x, y)$. Mais cette série peut-être calculée par la méthode DSV en résolvant une seule équation quadratique (et non un système de telles équations). On obtient ainsi aisément l'expression de $G(x, y)$.

1. Introduction

Self-avoiding polygons (*SAP*'s, Fig. 1) have a wide range of applications in physics and chemistry, and also have a strong appeal to pure combinatorialists. For the present, there are very little chances to compute the perimeter or area generating functions of general SAP's. However, for various special kinds of SAP's exact results are known. The definitions of those amenable subsets of SAP's often require convexity with respect to one or more directions.

This paper is focused on the remarkable class of *column-convex polyominoes* (*cc-polyominoes*). By definition, a cc-polyomino is such a SAP whose intersection with any vertical straight line is a convex set. See Fig. 2.

Two cc-polyominoes are considered as not different iff they can be transformed one into the other by a translation.

PII S0012-365X(97)00114-3

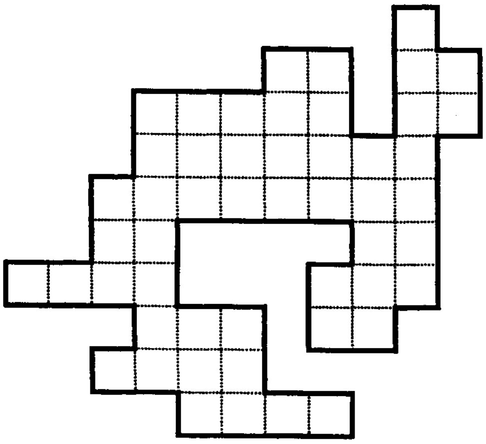

Fig. 1. A self-avoiding polygon (SAP).

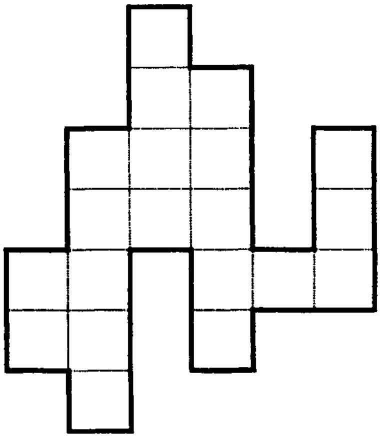

Fig. 2. A column-convex (cc-)polyomino.

Notation 1. Let P be a cc-polyomino. We shall write $h(P)$ for the number of horizontal edges of P, and $v(P)$ for the number of vertical edges of P.

Definition 1. Let Ω be some family of cc-polyominoes. By the perimeter generating function (gf) for Ω we mean the formal sum

$$\sum_{P \in \Omega} x^{h(P)} y^{v(P)}.$$

Column-convex polyominoes apparently first appeared in Pólya's 1938 diary notes [12], and were independently introduced by Temperley [14] in 1956. The area gf of this model was found on the spot [12, 14]. On the contrary, the perimeter gf of cc-polyominoes (and not to speak of their perimeter + area gf) remained unknown for many years after Temperley's and Pólya's work. At last Delest [5] applied the DSV-methodology [1,6,13,15,16] and the computer algebra program MACSYMA to obtain a formula for $G(x,x)$. Subsequently, Brak et al. [4] rederived the function $G(x,x)$ using the Temperley methodology and *Mathematica.* Thus it turned out that the formula given in [5] can be written in a simpler form. The result of Brak et al. [4] was generalized to the case $x \neq y$ by Lin [10].

In the course of preparation of their paper [7], Feretić and Svrtan were at first using the DSV-methodology. So they encoded the cc-polyominoes and set up a system of four nonlinear equations. Some manipulation of this system left them with a single degree-four algebraic equation satisfied by $G(x,y)$. Wishing to calculate the then unknown Taylor coefficients of $G(x,y)$ by use of the Lagrange inversion formula (LIF), they factored that algebraic equation. The result was that

$$H = \frac{x^2y^2(1-H)^4}{(1-y^2)^2(1-2H)[(1-3H)^2 - x^2(1-H)^2]}, \tag{1}$$

where

$$H = \frac{G}{1-y^2}. \tag{2}$$

After the application of the LIF to (1), an expression for $\langle x^{2c}y^{2v}\rangle G$ was readily found. (That expression is similar to the formula (24) of this paper). But after a while the authors of [7] divided Eq. (1) by $1-H$ and soon got a conspicuous bonus. Namely, it turned out that the function

$$L = (1-3H)/(1-H), \tag{3}$$

satisfies the biquadratic equation

$$L^4 - (1+x^2)L^2 + x^2\left(\frac{1+y^2}{1-y^2}\right)^2 = 0. \tag{4}$$

Solving the Eq. (4) and using $H = 1 - 2/(3-L)$, the following remarkably simple formula for $G(x,y)$ was obtained:

$$G(x,y) = (1-y^2)\left[1 - \frac{2\sqrt{2}}{3\sqrt{2} - \sqrt{1+x^2+\sqrt{(1-x^2)^2 - 16x^2y^2/(1-y^2)^2}}}\right]. \tag{5}$$

In [7] there is also an alternative proof for (5), which was found later. This second proof uses the Temperley recurrences, but these recurrences are solved by a method different from that in [4,10].

The aim of the present paper is to explain where do the apparently magic functions H and L come from. In Section 2, we introduce two new classes of plane figures, whose abbreviated names are tapoes and stapoes. Let I be the perimeter gf of the stapoes. We introduce the function $J = I/(1-y^2)$, and we readily establish the simple relationship $H = J/(1+J)$. In Section 3, the DSV-methodology is used to derive the function I and thereby the functions J, H and G. The computation is easy because there is just one quadratic equation to solve, whereas in [5,7] one had to handle systems of quadratic equations. The paper ends with some brief remarks, which are concerned mainly with the coefficients of G.

2. Two new objects

Let P be an arbitrary cc-polyomino. The top left corner of the first column of P is called the *northwest pole* of P and is denoted by $\mathrm{NW}(P)$. The bottom right corner of the last column of P is the *southeast pole* of P (notation: $\mathrm{SE}(P)$).

Imagine a plane figure T obtained by appending a vertical segment of $d \in \mathbf{N}_0$ lattice units to the southeast pole of a cc-polyomino P. We say that T is a *tailed polyomino* (a *tapo*, for short). Naturally, the appended segment is termed the *tail* of T. By the columns of a tapo T, we mean the columns of the underlying cc-polyomino P. The northwest pole of T is defined by $\mathrm{NW}(T) = \mathrm{NW}(P)$, while the southeast pole $\mathrm{SE}(T)$ is defined to be the lowest point of the tail of T, see Fig. 3.

Now let us define the second new object. Suppose that, for some $n \in \mathbf{N}$, $n-1$ arbitrary tapoes $T_1, \ldots, T_{n-1}$ and a tapo with a null tail T_n are given. Let $T_1, \ldots, T_n$ be disposed in a way that, for $2 \leqslant i \leqslant n$, the northwest pole of T_i coincides with the southeast pole of T_{i-1}.

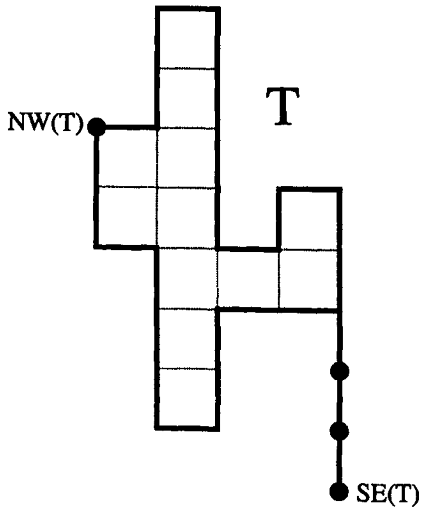

Fig. 3. A tailed polyomino (tapo) with 4 columns and 22 vertical edges.

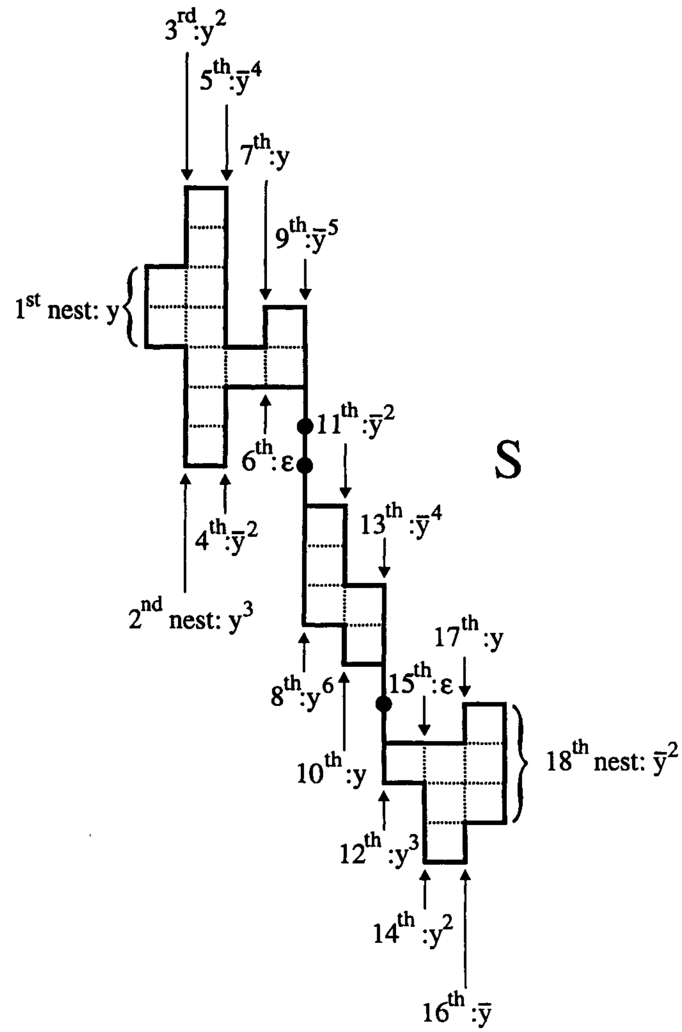

Fig. 4. A stapo and the nests of its code.

In a situation like this we say that the union $S = \bigcup_{1 \leqslant i \leqslant n} T_i$ is a *stapo* (short for: a *sequence of tailed polyominoes*). The tapoes $T_1, \ldots, T_n$ are called the *parts* of S. By the columns of a stapo we mean the columns of its parts. See Fig. 4. Observe that the one-part stapoes are cc-polyominoes.

It is useful to adopt the following convention.

Convention. Let a tapo T be obtained by appending a tail of length d to a cc-polyomino whose vertical perimeter is $2v$ (i.e. which has $2v$ vertical edges). Then T has $2v + 2d$ vertical edges.

By the vertical perimeter of a stapo we mean the sum of the vertical perimeters of its parts.

With this convention, in the sequel we shall apply Notation 1 and Definition 1 not only to the cc-polyominoes, but also to the tapoes and stapoes.

Let H_d be the perimeter gf for the tapoes whose tail is exactly d units long. It is obvious that $H_d = y^{2d}G$. By this remark and (2), the perimeter gf for all the tapoes is

$$\sum_{d\geqslant 0} y^{2d}G = \frac{G}{1-y^2} = H. \tag{6}$$

An n-part stapo is, in substance, a sequence of $n-1$ tapoes and one cc-polyomino. Hence, the perimeter gf for the n-part stapoes is $H^{n-1}G$. Let I be the perimeter gf for all the stapoes. We have

$$I = \sum_{n\geqslant 1} H^{n-1}G = \frac{G}{1-H}. \tag{7}$$

Further, it is convenient to put

$$J = I/(1-y^2). \tag{8}$$

The function J can be interpreted as the perimeter gf for the generalized stapoes, whose last part, too, is allowed to have a tail. From (2) and (7) it follows that

$$J = \frac{H}{1-H}, \tag{9}$$

so that

$$H = 1 - \frac{1}{1+J}. \tag{10}$$

3. The DSV-computation of the function G

3.1. Preliminaries on words and languages

Mostly due to the papers of the Bordeaux group for enumerative combinatorics [5–16], the algebraic language (i.e. DSV-) methodology is today a popular counting technique. Here we shall dispense with giving a complete introduction to this method. However, we shall give some mainly non-standard definitions concerning the free monoid $\{x, y, \bar{y}\}^*$, which is the one relevant to our forthcoming proof.

For $v \in \{x, y, \bar{y}\}^*$, we put $\delta(v) = |v|_y - |v|_{\bar{y}}$, and say that $\delta(v)$ is the *rank* of v.

Let $w \in \{x, y, \bar{y}\}^*$ and let $|w|_x = n$.

Clearly, w can be written as $u_1 \cdot x \cdot u_2 \cdot x \cdots u_n \cdot x \cdot u_{n+1}$, where $u_i \in \{y, \bar{y}\}^*$, for every i. The word u_i will be called the ith *nest* of w.

Then, we put $\delta_0(w)=0$, $\delta_1(w)=\delta(u_1)$ and

$$\delta_i(w)=\delta(u_1\cdot x\cdot u_2\cdots x\cdot u_i)\quad (i=2,\ldots,n+1).$$

When convenient, we shall omit the argument, and write δ_i instead of $\delta_i(w)$.

We say that w is a *Motzkin word* if the rank of w is zero and the ranks of all the left factors of w are non-negative.

We say that w is a word with *pure* nests if every nest of w belongs either to $\{y\}^*$ or to $\{\bar{y}\}^*$.

Also, we define $\tau(w)$ to be the word obtained from w by swapping the nests u_i and u_{i+1}, for every $i=2,4,\ldots,2\lfloor n/2\rfloor$.

For example, let $w=y\cdot x\cdot yy\cdot x\cdot x\cdot \bar{y}\cdot x\cdot y\cdot x\cdot \bar{y}\bar{y}\bar{y}$. This w is a Motzkin word and has pure nests. Further, we see that the numbers $\delta_0(w),\ldots,\delta_6(w)$ are $0,1,3,3,2,3,0$ and that $\tau(w)=y\cdot x\cdot x\cdot yy\cdot x\cdot y\cdot x\cdot \bar{y}\cdot x\cdot \bar{y}\bar{y}\bar{y}$.

The language formed by the Motzkin words with pure nests will be denoted $\mathscr{B}$.

3.2. A coding for the stapoes

For $c,v\in\mathbf{N}$, we shall write $\mathscr{S}_{cv}$ for the set of stapoes which have c columns and $2v$ vertical edges. When there is just one $S\in\mathscr{S}_{cv}$ under discussion, we write y_i and Y_i to mean the minimal and the maximal ordinate of the ith column of S.

Definition 2. The *code* of a stapo $S\in\mathscr{S}_{cv}$ is the word $w=\psi(S)$ which has the following properties:

(i) $w\in\{x,y,\bar{y}\}^*$;

(ii) w has pure nests;

(iii) $|w|_x=2c-1$;

(iv) $\delta_{2i-1}(w)=Y_i-y_i-1\quad (i\in\underline{c})$, [1]

$$\delta_{2i}(w)=Y_i-y_{i+1}-1\quad (i\in\underline{c-1}),\ \text{and}\ \delta(w)=0.$$

(It is clear that the above conditions define w uniquely.) Essentially, we encode the stapoes as Delest [5] encoded the cc-polyominoes. An example for our coding is shown in Figs. 4 and 5.

The task ahead of us is to characterize the language formed by the codes we have just defined. To do that, we need to make ourselves a little more familiar with the geometry of stapoes.

Remark 1. Every $S\in\mathscr{S}_{cv}$ has the properties $Y_i>y_i$ $(i\in\underline{c})$ and $Y_i>y_{i+1}$ $(i\in\underline{c-1})$. An $S\in\mathscr{S}_{cv}$ is a cc-polyomino if and only if it also has the property $Y_{i+1}>y_i$ $(i\in\underline{c-1})$.

[1] The symbol $\underline{c}$ denotes the set of integers $\{1,2,\ldots,c\}$.

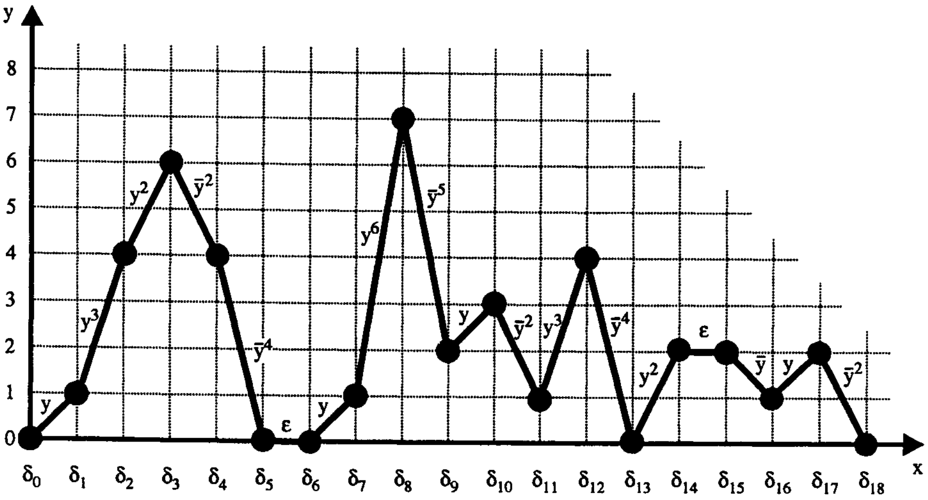

Fig. 5. The 'partial ranks' of the word which encodes the stapo of Fig. 4, i.e. the numbers $\delta_i(\psi(S))$, $i=0,1,\ldots,18$.

Remark 2. Let $S\in\mathscr{S}_{cv}$. The stapo S is incident on $c+1$ of the vertical lattice lines. Let these lines be $l_0,\ldots,l_c$, and let the contribution of $S\cap l_i$ to the vertical perimeter of S be denoted by v_i. Clearly,

$$v_0=Y_1-y_1 \quad\text{and}\quad v_c=Y_c-y_c.$$

Further, a little thought shows that for $i\in\underline{c-1}$ we have

$$v_i=|y_{i+1}-y_i|+|Y_{i+1}-Y_i|,$$

whether or not $S\cap l_i$ contains a tail of S. Therefore,

$$(Y_1-y_1)+(Y_c-y_c)+\sum_{i=1}^{c-1}(|y_{i+1}-y_i|+|Y_{i+1}-Y_i|)=v(S)=2v. \tag{11}$$

Let $S\in\mathscr{S}_{cv}$. Remark 1 tells us that the word $w=\psi(S)$ has the property $\delta_i(w)\geqslant 0$, $(\forall i\in\underline{2c-1})$. Since $\delta(w)=0$ and the nests of w are pure, this means that w is a Motzkin word. Using (11), we also find that

$$\begin{aligned}|w|_y+|w|_{\bar{y}}&=\delta_1+\delta_{2c-1}+\sum_{i=1}^{c-1}[|\delta_{2i}-\delta_{2i-1}|+|\delta_{2i+1}-\delta_{2i}|]\\&=(Y_1-y_1-1)+(Y_c-y_c-1)\\&\quad+\sum_{i=1}^{c-1}[|y_{i+1}-y_i|+|Y_{i+1}-Y_i|]=2v-2.\end{aligned} \tag{12}$$

Let

$$\mathscr{B}_{cv}=\{u\in\mathscr{B}\colon\ |u|_x=2c-1,\ \ |u|_y+|u|_{\bar{y}}=2v-2\}.$$

What we have proved so far is that our coding ψ maps the set $\mathscr{S}_{cv}$ into $\mathscr{B}_{cv}$. Now, observe that ψ preserves the $Y_i - y_i$ and $Y_i - y_{i+1}$ data, and that knowing these data, we have enough information to build up a stapo. This means that ψ is an injection, and it is a simple matter to see that ψ is also a surjection. Thus, we have:

Proposition 1. *ψ is a bijection between $\mathscr{S}_{cv}$ and $\mathscr{B}_{cv}$.*

Let $w \in \mathscr{B}_{cv}$ and let S be the stapo encoded by w. Under what condition on w is S a cc-polyomino? To answer this question, we take a look at the word $\tau(w)$.

First, $\tau(w)$ obviously has pure nests as well as the properties

$$|\tau(w)|_x = 2c-1, \quad |\tau(w)|_y + |\tau(w)|_{\bar{y}} = 2v-2 \quad \text{and} \quad \delta(\tau(w)) = 0.$$

Second, Definition 2(iv) implies that

$$\delta_{2i-1}(\tau(w)) = Y_i - y_i - 1 \quad (i \in \underline{c}) \tag{13}$$

and

$$\delta_{2i}(\tau(w)) = Y_{i+1} - y_i - 1 \quad (i \in \underline{c-1}). \tag{14}$$

Now, if S is a cc-polyomino, Remark 1 tells us that $\delta_j(\tau(w)) \geqslant 0$, for every $j \in \underline{2c-1}$, and it follows that $\tau(w)$ lies in $\mathscr{B}_{cv}$. Conversely, if $\tau(w)$ lies in $\mathscr{B}_{cv}$, then the δ_{2i}'s of $\tau(w)$ are all non-negative, so that $Y_{i+1} > y_i$, for every $i \in \underline{c-1}$. By Remark 1, this means that S is a cc-polyomino.

Let

$$\mathscr{P}_{cv} = \{P \in \mathscr{S}_{cv}\colon P \text{ is a cc-polyomino}\}$$

and let

$$\mathscr{B}_{cv}^{+} = \{w \in \mathscr{B}_{cv}\colon \tau(w) \text{ also lies in } \mathscr{B}_{cv}\}.$$

To sum up, our conclusion is:

Proposition 2. *ψ induces a bijection between $\mathscr{P}_{cv}$ and $\mathscr{B}_{cv}^{+}$.*

Now, the absence of the awkward requirement '$\tau(w) \in \mathscr{B}_{cv}$' indicates that it will probably be easier to enumerate the family $\mathscr{B}_{cv}$ than the family $\mathscr{B}_{cv}^{+}$. In fact, it was right for this reason that stapoes have been introduced.

3.3. The grammar and the algebra

We define a power series $B(x,y)$ as follows. For $i,j \in \mathbf{N}_0$, the coefficient of $x^i y^j$ in B, usually written as $\langle x^i y^j \rangle B$, is

$$\operatorname{card}\{w \in \mathscr{B}\colon |w|_x = i, \ |w|_y + |w|_{\bar{y}} = j\}.$$

Next, let $D(x,y)=[B(x,y)-B(-x,y)]/2$. The definitions and Proposition 1 imply that $\langle x^{2c}y^{2v}\rangle I$ equals $\langle x^{2c}y^{2v}\rangle xy^2D$ for all $c,v\in\mathbf{N}$. Since in the power series I and xy^2D all the powers of x and y are even, this means that

$$I=xy^2D. \tag{15}$$

It is readily seen that the language $\mathscr{B}$ is generated by the unambiguous grammar

$$\mathscr{B}=\varepsilon+x\mathscr{B}+y(\mathscr{B}-\varepsilon)\bar{y}(\varepsilon+x\mathscr{B}). \tag{16}$$

Letting the letters in (16) commute and putting $y=\bar{y}$, we find that the gf $B(x,y)$ satisfies the quadratic equation

$$B=1+xB+y^2(B-1)(1+xB),$$

which is equivalent to

$$xy^2B^2-(1-x)(1-y^2)B+1-y^2=0. \tag{17}$$

Solving Eq. (17) we obtain

$$B=\frac{(1-x)(1-y^2)-\triangle_-^{1/2}}{2xy^2}, \tag{18}$$

and

$$D=\frac{2(1-y^2)-\triangle_-^{1/2}-\triangle_+^{1/2}}{4xy^2}, \tag{19}$$

where

$$\begin{aligned}\triangle_- &=(1-x)^2(1-y^2)^2-4xy^2(1-y^2),\\ \triangle_+ &=(1+x)^2(1-y^2)^2+4xy^2(1-y^2).\end{aligned} \tag{20}$$

Using (15) to obtain I, (8) to obtain J, (10) to obtain H and (2) to obtain G, we have the following theorem:

Theorem 1. *The perimeter generating function for the column-convex polyominoes is given by*

$$G(x,y)=(1-y^2)\times\left[1-\frac{4}{6-\sqrt{(1-x)^2-\frac{4xy^2}{(1-y^2)}}-\sqrt{(1+x)^2+\frac{4xy^2}{(1-y^2)}}}\right]. \tag{21}$$

4. Final remarks

Let $\delta_-^{1/2}$ and $\delta_+^{1/2}$ denote the first and the second of the square roots which appear in the denominator of (21).

Contrary to appearances, (21) is equivalent to the formula for G quoted in the Introduction. This fact is related to the possibility to write $\delta_-^{1/2}+\delta_+^{1/2}$ as $[\delta_- + \delta_+ + 2(\delta_-\delta_+)^{1/2}]^{1/2}$.

Once we have found the formula for G, the biquadratic Eq. (4) can be established with ease. Indeed, we just need to rewrite (21) as

$$\delta_-^{1/2} = 2L - \delta_+^{1/2}, \tag{22}$$

(where L is given by (3)), and then suitably square two times.

If we want to compute the coeficients of G, we may start by rearranging Eq. (4) into

$$1 - L = \frac{4x^2y^2}{(1-y^2)^2(L+1)(L^2-x^2)}. \tag{23}$$

Plugging $L = 1 - 2J$ (which follows from (3) and (9)) into (23), we obtain an algebraic equation for J. An application of the Lagrange inversion formula to that latter equation produces the power series expansion of J^r for general $r \in \mathbf{N}$. Combining that expansion with the relations $H = J/(1+J)$ and $G = (1-y^2)H$, we eventually have[2]

$$\begin{aligned}\langle x^{2c}y^{2v}\rangle G = \frac{1}{c}\sum_{i,j,k\geqslant 0} (-1)^k(k+1)\binom{v+i-1}{2i}\binom{c}{i+1}\\ \times \binom{2c+j-1}{j}\binom{2c+2i-k}{i-j-k}.\end{aligned} \tag{24}$$

Certainly, we would be more content if the right-hand side of (24) were in closed form. To all appearance, however, here we have nothing left but to settle for threefold sums.

Besides the horizontal and vertical perimeters, the stapo-method can take into account also the heights of the first and last columns as well as some other properties of a cc-polyomino. In this paper, however, we confined our attention to the two perimeters, because we have taken the view that the method should be presented in its simplest form.

References

[1] M. Bousquet-Mélou, Convex polyominoes and algebraic languages, J. Phys. A 25 (1992) 1935–1944.

[2] M. Bousquet-Mélou, A method for the enumeration of various classes of column-convex polygons, Discrete Math. 154 (1996) 1–25.

[3] R. Brak, A.J. Guttmann, Exact solution of the staircase and row-convex polygon perimeter and area generating function, J. Phys. A 23 (1990) 4581–4588.

[4] R. Brak, A.J. Guttmann, I.G. Enting, Exact solution of the row-convex polygon perimeter generating function, J. Phys. A 23 (1990) 2319–2326.

[5] M.P. Delest, Generating functions for column-convex polyominoes, J. Combin. Theory Ser. A 48 (1988) 12–31.

[2] In (24) we assume that $\binom{a}{b}$ is zero whenever $a<0$ or $b<0$ (or both).

[6] M.P. Delest, G. Viennot, Algebraic languages and polyominoes enumeration, Theoret. Comput. Sci. 34 (1984) 169–206.
[7] S. Feretić, D. Svrtan, On the number of column-convex polyominoes with given perimeter and number of columns, in: A. Barlotti, M. Delest, R. Pinzani (Eds.), 5th FPSAC Proc., Firenze, 1993, pp. 201–214.
[8] I.M. Gessel, A combinatorial proof of the multivariable Lagrange inversion formula, J. Combin. Theory Ser. A 45 (1987) 178–195.
[9] D.A. Klarner, Some results concerning polyominoes, Fibonacci Quart. 3 (1965) 9–20.
[10] K.Y. Lin, Perimeter generating function for row-convex polygons on the rectangular lattice, J. Phys. A 23 (1990) 4703–4705.
[11] K.Y. Lin, W.J. Tzeng, Perimeter and area generating functions of the staircase and row-convex polygons on the rectangular lattice, Internat. J. Mod. Phys. B 5 (1991) 1913–1925.
[12] G. Pólya, On the number of certain lattice polygons, J. Combin. Theory 6 (1969) 102–105.
[13] M.P. Schützenberger, Context-free languages and pushdown automata, Inform. Control 6 (1963) 246–264.
[14] H.N.V. Temperley, Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules, Phys. Rev. 103 (1956) 1–16.
[15] G. Viennot, Enumerative combinatorics and algebraic languages, in: L. Budach (Ed.), Proc. FCT'85, Lecture Notes in Computer Science, vol. 199, Springer, Berlin, 1985, pp. 450–464.
[16] X.G. Viennot, A survey of polyominoes enumeration, in: P. Leroux, C. Reutenauer (Eds.), 4th FPSAC Proc., Publications du LACIM, vol. 11, Montréal, 1992, pp. 399–420.